Probability and Statistics: Assignment 1.

1. The excel sheet 'gender-height-weight' contains information about the height and weight of males and females who participated in a program about the general fitness. The given indexes have the following meaning:

Index	Body type
0	Extremely Weak
1	Weak
2	Normal
3	Overweight
4	Obesity
5	Extreme Obesity

   a) Using the excel sheet, you have to complete the following tasks:
   b) Distinguish the data on the basis of numeric, categorical and nominal
   c) Bring the file into the pandas dataframe.
   d) Using appropriate functions display:
      i.   the percentage of males and females in this data set
      ii.   the percentage of obese male and females in the group
      iii.   the attributes used in the data set
      iv.   the number of observations
      v.   size of the data set
   e) Plot two histograms side-by-side, showing the heights of males and females in the data-set.
   f) Calculate the average height and weight of males and females who participated in the study
   g) Find the range, sample variance, and sample standard deviation for the males and females participants. Explain what results you get
   h) Find the inter quartile range.
   i) Are there any outliers in the data?
   j) Make a box plot for weight vs. gender. What information you get from it?
   k) Make a histogram by grouping the data with interval 10 start from 50 (Weight)
   l) What statistics can you use to infer that obesity does/ does not depend on gender
   m) Which Python libraries did you use for completing this task? Was there any additional function that you used but it was not discussed in class ?